European Journal of Orthodontics, 2019, 1–12
doi:10.1093/ejo/cjz075
Systematic Review

Systematic Review

Assessment of techniques used for superimposition of maxillary and mandibular 3D surface models to evaluate tooth movement: a systematic review

Sven Stucki and Nikolaos Gkantidis

Department of Orthodontics and Dentofacial Orthopedics, University of Bern, Switzerland

Correspondence to: Nikolaos Gkantidis, Department of Orthodontics and Dentofacial Orthopedics, University of Bern, Freiburgstrasse 7, CH-3010 Bern, Switzerland. E-mail: nikolaos.gkantidis@zmk.unibe.ch; nikosgant@yahoo.gr

Summary

Background: Superimposition of three-dimensional (3D) digital models offers great opportunities to assess tooth movement during time. In the literature, several superimposition techniques are described.
Objectives: To summarize and critically assess the available evidence from studies on serial digital 3D dental model superimposition.
Search methods: MEDLINE *via* Ovid and PubMed, EMBASE *via* Ovid, Cochrane Register of Diagnostic Test Accuracy Studies, and Google Scholar were searched with no time limit (last update: December 2018). Hand and unpublished literature searches were also performed.
Selection criteria: Studies of any design that had a sample size ≥5 and tested superimposition of serial 3D digital dental models to assess tooth movement were included.
Data collection and analysis: Study selection, data extraction, and risk of bias assessment were performed independently by the authors.
Results: Twelve studies met the eligibility criteria. The total risk of bias (QUADAS-2 tool) of 10 studies was high, whereas only 2 studies had unclear bias. Ten studies had high and two studies low overall applicability concerns. From these, one study tested the mandibular alveolar bone area, three studies the rugae area, nine studies a larger palatal area, and two studies the incisive papilla area. The high heterogeneity in samples, outcomes, and methods did not allow for synthesis of a considerable amount of studies in any case.
Limitations: The high heterogeneity among studies and the limited evidence did not allow for solid conclusions.
Conclusions and implications: The following areas of the maxilla could provide reliable outcomes: (1) the medial two-thirds of the third rugae and the area 5 mm dorsal to them and (2) an area including all rugae, with the lateral margins located at least 5 mm from the gingival margins and a distal margin that does not extend beyond the first molars. No recommendation is possible for other regions of the mouth. There is an urgent need for further research in the field.
Registration: PROSPERO (CRD42019124365).

Introduction

Since many years, researchers and clinicians that work in the craniofacial area use superimpositions of serial patient images as a mean to assess treatment outcome and subsequently generate data that would allow valid predictions of treatment effects in future patients. Traditionally, superimpositions to assess tooth movement have been made through 2D (2-dimensional) cephalometric image superimpositions, though these have various inherent limitations (1).

With recent technological advances, further possibilities to superimpose serial 3-dimensional (3D) patient images became available. Nowadays, plaster models or directly the oral structures can be scanned with appropriate scanners to create detailed digital 3D representations of a patient's mouth. The digital dental models became quite popular in the last years due to various advantages such as real size detailed information, no need for physical storage, risk- and cost-free transfer, and extensive potential for data processing (2). To allow for digital superimposition of two or more such 3D models several computer software programs and techniques have been introduced in recent years (2). Compared to the superimposition of lateral cephalometric radiographs this new possibility offers various advantages, such as no tracing errors, real size representation of structures (no magnification), no image distortion, no radiation, and no dimensional reduction (3D information) (1, 3). The cone-beam computed tomography is the radiographic equivalent of a lateral cephalogram that offers 3D information, but radiation exposure is still required and it is even higher than that of a lateral cephalometric radiograph (4). Furthermore, the quality of the obtained digital 3D dental models is much lower compared to that from current intraoral scanners (5).

In recent literature, different methods for the implementation of serial 3D dental model superimpositions have been described, such as landmark-based (6–9) or surface-based approaches (2, 3, 6, 7, 9–15). Landmark-based superimposition requires manual identification of a certain number of corresponding anatomical landmarks in both models that are going to be registered, whereas surface-based approaches usually require a reference area selection only in one model (2). There are various options for landmark-based superimposition depending on the location and number of points selected and the technique used to superimpose them. Similarly, different palatal regions have been suggested as references for surface-based superimposition. It has recently been shown that the reference area used has a considerable effect on the superimposition outcome (2), as well as the number of landmarks used on the landmark-based approaches (6). So far, there is no consensus in the literature regarding the techniques to superimpose serial 3D intraoral digital models.

Thus, the aim of the present study is to summarize and critically assess the available evidence from studies on serial digital 3D dental model superimposition. We will evaluate techniques concerning the maxilla and the mandible aiming to provide to both clinicians and researchers a critical overview of the available knowledge, in order to aid them in implementing these techniques in everyday practice and form the basis for future research.

Materials and methods

Protocol and registration

The protocol was registered in PROSPERO prior to the study (Registration No. CRD42019124365).

Search strategy

To identify eligible studies specific search strategies were developed for the following databases to fulfil their individual controlled vocabulary and syntax rules: MEDLINE *via* Ovid and PubMed, EMBASE *via* Ovid, Cochrane Register of Diagnostic Test Accuracy Studies, Google Scholar. Studies published at any time till 10 December 2018 were evaluated. Hand searches of the eligible studies and systematic or narrative reviews on relevant topics were also performed. Unpublished literature was searched through the National Research Register, and Pro-Quest Dissertation Abstracts and Thesis database. The exact search is shown as supplementary information (Supplementary Table 1).

Selection criteria applied for the review

1. Study design: Any study design was considered eligible, including prospective, and retrospective studies of any type.
2. Study sample: Studies with sample size ≥5.
3. Index test: Surface-based or landmark-based superimposition of serial 3D digital dental models.
4. Types of participants: Patients who received any kind of orthodontic treatment or whose dental models were used to simulate treatment.
5. Type of intervention: 3D superimposition to assess any orthodontic or simulated tooth movement.
6. Primary outcome: Accuracy or precision of a superimposition technique, or agreement between different techniques measured as tooth movement or area distance between corresponding models.

Studies that evaluated any of the aforementioned parameters as a secondary outcome were also included.

7. Comparator/control group: Different superimposition techniques, direct measurements, or repeated measurements.
8. Unit of analysis: In all cases, the unit of analysis was the tooth or the measured area.
9. Follow-up: All observation periods between subsequent models were accepted.
10. Exclusion criteria: Non-human-derived data.

Study selection

The selected databases were searched by the two authors of the review. They were not blinded to the identity of the authors of the studies, their institutions, or the results of their research. The studies were selected by reading the title as well as the abstract and if necessary the full text. Non-eligible articles were excluded. The left over studies were read again in full text and eligible studies were included independently by the authors of the review. If there was a disagreement, the eligibility was discussed between the authors until a consensus was reached. A record of all decisions on study identification was kept.

Data extraction

Data extraction was performed independently and in duplicate by the two authors. The following information was extracted from all eligible studies, if available:

1. Methods: Author, title, year, objectives, and design of study.
2. Participants: Number, age, gender of patients recruited.
3. Materials: Type of 3D model acquisition method and time between serial models.
4. Superimposition method: Type of superimposition reference areas or points and software used.
5. Comparison/control group: Type and characteristics.
6. Outcome: Type of outcome(s) and method of outcome assessment.

When needed, by the presence of missing data in a study the authors were contacted by email to request the information. In case that the authors did not respond or the data were not receivable, only the available information was regarded.

Assessment of heterogeneity

We assessed heterogeneity by examining the characteristics of the studies, the similarity between the types of participants, the methods compared, and the outcomes as specified in the inclusion criteria.

Assessment of reporting bias

Reporting biases arise when the reporting of research findings is affected by the nature or direction of the findings. We attempted to minimize potential reporting biases including publication bias and multiple (duplicate reports) publication bias, by conducting an accurate and at the same time a sensitive search of multiple sources. We also searched for on-going studies.

Data synthesis

We planned to conduct meta-analysis if there were at least two unclear or low risk of bias studies of similar comparisons, reporting the same outcomes at similar follow-up periods.

Subgroup analysis

Whenever possible, results will be assessed also in the following subgroups:

1. Extraction versus non-extraction treatment.
2. Patients with versus without growth.
3. Short-term (within 1 year) versus medium/long-term (>1 year) interval between serial casts.
4. Surface- versus landmark-based techniques.

Quality assessment

The QUADAS-2 tool (16) was used to evaluate the quality of the selected studies. This is a Cochrane Collaboration recommended tool to assess the risk of bias and the applicability concerns of diagnostic accuracy studies in systematic reviews. This tool subdivides each study into four key domains: patient selection, index test, reference standard, and flow and timing. Those four domains are evaluated in two categories, namely the risk of bias and the applicability concerns. The results of the QUADAS-2 tool are commonly presented in a table using happy (low risk) or sad smiles (high risk). If it is not possible to evaluate a domain an interrogation mark is set, which means an unclear risk. In case a domain is evaluated with a high risk an explanation is provided.

Risk of bias was performed independently and in duplicate by the two authors. In cases of disagreement the judgment was discussed between the authors until a consensus was reached. In case of a meta-analysis, studies with high risk of bias would have not been included.

Results

Description of studies

The flow diagram (17) is shown in Figure 1. We identified 690 studies through database searching. Twelve studies were added through hand searches. We eliminated duplicates and received a total of 312 studies that we screened through title and abstract reading. Thirty-one studies had to be read in full text to assess eligibility. Eleven studies were excluded as irrelevant and eight studies were excluded with an argumentation shown in the Supplementary Text. Finally, 12 articles were included in this review.

In all studies, 3D superimposition techniques were used to assess an orthodontic or simulated tooth movement. All studies evaluated

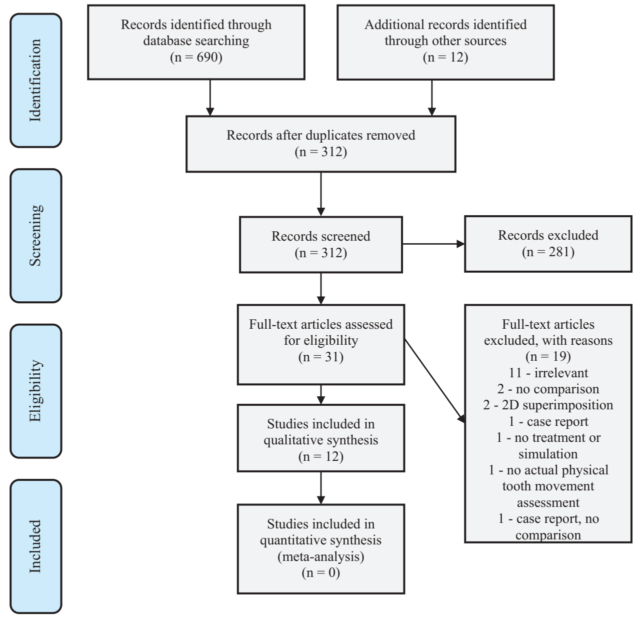

Figure 1. Flowchart of study selection according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.

the accuracy or precision of a superimposition technique, or agreement between different techniques measured as tooth movement or area distance between corresponding models, as primary or secondary outcome.

Quality assessment

The quality assessment using the QUADAS-2 tool is shown in Supplementary Table 2. There is an explanation included if a study is graded with a high risk or unclear risk of bias, as well as if a study has high or unclear applicability concerns.

The total risk of bias of 10 of the 12 included studies was high, whereas only 2 studies had an unclear risk. In the individual items, all 12 included studies had a low risk regarding the patient domain. Regarding the index test, nine studies had a high risk of bias, two were graded with low risk of bias, and one study with an unclear risk of bias. The reference standard of six studies had a high risk of bias, of three studies a low risk, and of three studies an unclear risk of bias. The flow and timing of 11 studies had a low risk of bias and of 1 study had an unclear risk of bias.

Concerning the applicability concerns, 10 of the 12 included studies had high and 2 low overall applicability concerns. In the individual items, six studies had low and six studies had high applicability concerns regarding the patient selection. 10 studies had high and 2 studies low applicability concerns regarding the index test. The reference standard of four studies was of high applicability concerns, whereas six studies had low, and two unclear applicability concerns.

Characteristics of the included studies

From the 12 included studies, 7 studies were retrospective, 4 studies were prospective, and 1 had a cross-sectional design in terms of material origin. However, all studies were prospective in terms of data generation and method comparison. In three studies the patients were non-growing, four studies regarded only growing patients, three studies had mixed samples, and two studies did not mention patient age.

Regarding the superimposition reference area, there is only one study that used the alveolar bone. Three studies superimposed the 3D models on the rugae region. Superimpositions on a major region of the palate were used in nine studies. Finally, two studies superimposed on or adjacent to the incisive papilla area.

The main general and superimposition related characteristics of the included studies are shown in Table 1. More detailed information is provided in Supplementary Tables 3 and 4.

Results and qualitative synthesis of included studies

The heterogeneity of the studies regarding study design, samples, techniques, and measured outcomes was high. Therefore, a quantitative synthesis of the results was not possible.

The main analysis related characteristics and conclusions of the included individual studies are shown in Table 2. More detailed information is provided in Supplementary Tables 5 and 6.

For qualitative synthesis, the included studies will be subdivided in the four following categories: superimposition on the alveolar bone, superimposition on the rugae region, superimposition on a major region of the palate, and superimposition on or adjacent to the incisive papilla.

Superimposition on the alveolar bone

At the moment there is only one study that analysed the superimposition on alveolar bone structures and this regarded only the mandible (3). Using a surface-based superimposition method, the study tested the agreement of digital model superimposition on four different alveolar bone areas with cephalometric measurements. The study concluded that in patients with mandibular tori the superimposition outcome was more reliable. However, the study has important limitations, such as the too small subgroup sample size ($n = 5$) and the comparison with cephalometric radiographs. The study has a high risk of bias, as well as high applicability concerns, and it was the only one on the topic. Thus, further research is required to draw any conclusions.

Superimposition on rugae regions

There are three studies that specifically assessed the superimposition on the rugae region (2, 6, 11). The heterogeneity in this category is high, because all studies evaluated different superimposition areas. One study included growing patients (2), one non-growing, (11) and one both (6). The study of Chen *et al.* (11) searched for stable palatal surfaces to superimpose digital 3D models, using miniscrews as reference structures. According to this evaluation, the medial two-thirds of the third rugae and the area distal to it extending until the distal end of first molars remains stable within 0.5 mm. However, the stability of the screws was evaluated through linear measurements (0.5 mm allowance), the deviation between structures was only visually assessed, the accepted range was relatively large, and only mean comparisons between measurements were performed. The study has a high risk of bias, as well as high applicability concerns, and thus the results should be treated with caution.

Vasilakos *et al.* (2) evaluated the accuracy and precision of five palatal areas, previously used for surface-based superimposition of maxillary dental casts. They found that two areas, including the middle two-thirds of the third rugae show adequate accuracy, whereas the other three do not. The main limitation of this study is that the comparison is done against an assumed gold standard. Thus, the study has an unclear risk of bias, but low applicability concerns.

The study of Becker *et al.*(6) evaluated the agreement between two matching approaches on the assessment of orthodontic treatment outcome. The ICP (Iterative Closest Point) approach is a surface-based superimposition method, in which the digital 3D models are superimposed on a major region of the palate. The CP (control point)-based approach is a landmark-based superimposition method in which the digital models were superimposed on 10 landmarks located in the incisive papilla and the rugae region. The study states that the CP-based as well as the ICP-based superimposition method may both produce comparable results. Nevertheless, the study tests only mean values, the age range is large, and there is no method error. The study has a high risk of bias, as well as high applicability concerns.

In summary, though two of the three studies in this category showed promising results for the area of the middle two-thirds of the third rugae (2, 11), there was high heterogeneity in terms of samples and methods, and thus further research is required to validate and generalize this finding. The third study (6) assessed different areas and techniques, but due to its limitations the results are questionable.

Superimposition on a major region of the palate

There are nine studies that assess the superimposition on a major region of the palate. Three studies were done on growing patients (2, 9, 13), one on non-growing patients (7), three studies included growing as well as non-growing patients, (6, 10, 15) and in two studies the age was not reported (12, 14).

Four of the nine studies evaluated the same superimposition area (2, 12, 13, 15). This included all rugae, with the lateral margins located 5 mm from the gingival margins and a distal margin that does not extend distally beyond the distal surfaces of the maxillary first (2, 13, 15) or second molars (12). The incisive papilla was excluded. The study of Vasilakos *et al.* (2), which was previously reported, did not support the use of this area, whereas the other three studies did, and this was the only study that compared different areas. The study of Choi *et al.* (13) is a study on clinical data that confirmed the previous findings of an experimental study. The main limitation of Choi *et al.* (12) is that it is an *in vitro* study that simulated treatment to create pre- and post-treatment models, but it did not change the palatal vault. The study has a high risk of bias, as well as high applicability concerns. The study of Choi *et al.* (13) compared the digital model superimposition results with cephalometric radiographs and has an unclear risk of bias and unclear applicability concerns. The study of Yun *et al.* (15) had no comparison technique, evaluated only mean values, and had a small sample size. Furthermore, the study has a high risk of bias, as well as high applicability concerns.

Overall, one study with unclear risk of bias and low applicability concerns did not support this area (2), whereas two studies that suggested this area had high risk of bias and applicability concerns (12, 15) and one study had unclear risk of bias and low applicability concerns (13). Furthermore, the heterogeneity among studies was also high. To clarify this issue, these results require validation by different groups, with adequate methods, and in different samples.

The whole palate was tested as a superimposition reference by three other studies (9, 10, 14), which extended their lateral limits closer to the gingival margins of the posterior teeth (~1–3 mm distance). Otherwise, the reference area also included all rugae and extended posteriorly till the distal margins of the second premolars (14) or the first permanent molars (10). Cha *et al.* (10) compared surface-based 3D dental model superimposition with cephalometric

Table 1. Main general and superimposition-related characteristics of the included studies.

Study name	Objectives	Sample size and sex	Age	Growth	Type of participants (serial casts)	Time between serial casts	Superimposition Method	References for superimposition
An et al J Orofac Orthop (2015) (3)	Identification of stable reference areas for superimposing 3D dental models in the mandible.	n = 10 (4M, 6F; 5 with and 5 without mandibular torus)	Mean: 24.9±10 (range: 16.6–47.8) years	Non-growing	Pre- and post-orthodontic treatment: extraction of 4 premolars; patients with and without mandibular torus.	Mean: 33 months	Dental models: surface-based Lateral cephalograms: manually	Dental models: 1. the bilateral lingual surfaces of the alveolar bone of the premolar and molar area, 2. the lingual alveolar surface of the entire dentition, 3. the bilateral buccal and the lingual alveolar surfaces of the premolar and molar area, 4. the bilateral mandibular tori. Lateral cephalograms: lower border of the mandible and symphysis.
Becker et al J Orofac Orthop (2018) (6)	Evaluation of the agreement between two different matching approaches [control point (CP)- based vs. iterative closest point (ICP) matching] on the assessment of orthodontic treatment outcome.	n = 48 (22M, 26F)	Range: 11–53 years	Growing and non-growing	Orthodontic treatment: patients with an initial indication for upper molar protraction, who completed treatment with an appliance coupled with two mini-screws inserted into the anterior palate.	Mean: 11.65 ± 7.55 months	Dental models (CP superimposition): landmark-based Dental models (ICP superimposition): surface-based	Dental models with landmark-based matching: 10 landmarks located in the incisive papilla and rugae region. Dental models with surface superimposition: borders for ICP delimited by the connecting outline of the following 5 landmarks: incisal papilla, right and left gingival margin of the first premolar and right and left gingival margin of the second molars.
Cha et al Eur J Orthod (2007) (10)	Comparison of three-dimensional digital models superimposition with cephalometric superimposition.	n = 30 (6M, 24F)	Mean: 17.7 (range: 11.1–29.8) years	Growing and non-growing	Pre and post orthodontic treatment: orthodontic treatment with extraction of 4 premolars.	Mean: 35.3 (range: 26–51) months	Dental models: surface-based Lateral cephalograms: manual	Dental models: all rugae, lateral limits located approximately 1–3 mm from the gingival margin of the posterior teeth, and posterior limit at the distal margins of the first permanent molars. Lateral cephalograms: palatal plane registered at ANS.
Chen et al Orthod Craniofac Res (2011) (11)	Identification of a stable palatal region to superimpose serial maxillary dental models in adult extraction cases.	n = 15 (11M, 4F)	Mean: 25.8 (range: 21–41) years	Non-growing	Pre and during orthodontic treatment: extraction of maxillary first premolars to reduce protrusion; placement of 6 miniscrews (2 loaded, 4 unloaded); en masse retraction of anterior teeth.	Mean: 17 months (completion of the retraction)	Superimposition on minimum 3 (max 4) unloaded mini-screws to evaluate stable regions in the palate (< 0.5 mm distance). Superimposition on the identified stable region (PVR) and comparison of tooth movement results with the miniscrew superimposition.	Dental models: 1. superimposition on three or four stable unloaded miniscrews, 2. superimposition on the following stable region (region with a deviation < 0.5mm): medial 2/3 of the third rugae and the area distal to it extending until the distal end of first molars (3D-palatal-vault-regional-technique: PVR).

Table 1. Continued

Study name	Objectives	Sample size and sex	Age	Growth	Type of participants (serial casts)	Time between serial casts	Superimposition Method	References for superimposition
Choi et al Angle Orthod (2010) (12)	Evaluation of the superimposition of digital 3D models using the palate surface as reference for measuring tooth movements as compared to direct cast measurements.	n = 20 (gender NA)	NA	NA	Simulated treatment: random movement using a wax set up.	No second dental models (simulation)	Dental models: surface-based	Dental models: the palate including all rugae; the lateral margins were located at least 5mm from the gingival margins of the posterior teeth bilaterally, the distal margin did not extend distally beyond the line in contact with the distal surfaces of the maxillary second molars bilaterally, the incisive papilla was excluded.
Choi et al Korean J Orthod (2012) (13)	To assess the validity of 3D digital model superimposition in the palate in patients treated with rapid maxillary expansion (RME) and maxillary protration headgear.	n = 30 (12M, 18F)	Mean: 9.6±1.4 (range: 7.3–11.8) years	Growing	Pre and post orthodontic treatment: RME and maxillary protraction headgear treatment.	Mean: 8.4 ± 2.5 (range: 4.0–13.0) months	Dental models: surface-based on a palatal area Lateral cephalograms: manual	Dental models: area including the palatal rugae and palatal slope separated by 5 mm from the gingival margins of the bilateral posterior teeth, and not extending distally beyond the line in contact with the distal surfaces of the bilateral first molars. Lateral cephalograms: on the palatal plane with anterior nasal spine (ANS) as the registration point.
Ganzer et al Eur J Orthod (2017) (14)	Evaluation of a superimposition technique (named RFD), based on simulated tooth movement and growth.	n = 16 (gender not mentioned)	NA	NA	Simulated treatment: simulation of space closure after extraction of the first premolars and of growth by morphological change of the palatal vault.	No second dental model (initial model artificially altered)	Dental models: surface-based in three steps	Dental models: all rugae, lateral limits located approximately 1–3 mm from the gingival margin of the posterior teeth, and posterior limit at the distal margins of the second premolars.
Jang et al Angle Orthod (2009) (7)	Evaluation of a superimposition method for maxillary digital 3D models.	n = 10 (4M, 6F)	Mean: 20 (range: 15.6–27) years	Non-growing	Orthodontic treatment of maxillary protrusion including bilateral extraction of maxillary first premolars and placement of three miniscrews, ligated with a transpalatal arch to reinforce orthodontic anchorage.	NA	Dental models: surface-based after landmark-based	Dental models: 1. Point: the midpoint on the line connecting the medial points of the right and left third palatal rugae. Surface: the surface of the palatal vault surrounded by two transverse and two anteroposterior lines. One of the transverse lines is 10 mm away distally from the third palatal rugae, and the other is 5 mm away mesially from the line in contact with the distal surfaces of the maxillary second molars. On the palatal side, the anteroposterior lines are 10 mm from the lines in contact with the palatal gingival margins of the posterior teeth bilaterally. 2. Best fit on 3 loaded miniscrews. Two were placed on the palatal slopes bilaterally between the second premolar and the first molar. Another miniscrew placed on the paramedian region of the hard palate.

Table 1. Continued

Study name	Objectives	Sample size and sex	Age	Growth	Type of participants (serial casts)	Time between serial casts	Superimposition Method	References for superimposition
Nalcaci et al Korean J Orthod (2015) (8)	Evaluation of the reliability of measurements obtained after the superimposition of digital 3D models by comparing them with those obtained from lateral cephalomertric radiographs and photocopies of plaster models for the evaluation of upper molar distalization.	n = 20 (10M, 10F)	Mean: 16 years	Growing	Orthodontic treatment: distalization of the first molars with an intraoral distalizer for Class II correction.	NA	Dental models: landmark-based Lateral cephalograms: manual	Dental models: three points of the incisive papilla area (the most anterior point, the most prominent point, the most posterior point). Lateral cephalograms: a line perpendicular to the sella-nasion (SN) plane at the intersection of the anterior wall of the sella turcica and the anterior clinoid process.
Talaat et al Eur J Orthod (2017) (9)	Evaluation of the validity of 3D landmark-based palatal superimposition of digital models using Ortho Mechanics Sequential Analyser (OMSA) and comparison to the surface-based 3dMD superimposition of 3D dental models and the surface-based Invivo Dental superimposition of CBCTs.	n = 20 (gender NA)	Mean: 12.3 ± 1.9 (range: 8–15) years	Growing	Orthodontic treatment: maxillary expansion through Hyrax palatal expanders.	3 months	Dental models: Landmark-based or surface-based CBCT: surface-based	Dental models: three points (1. distal end of the incisive papilla, 2. and 3. arbitrary distal to the first point along the middle palatal suture) or palatal surface. CBCT: cranial base.
Vasilakos et al Sci Rep (2017) (2)	Evaluation of the accuracy and precision of the palatal areas, previously used for superimposition of maxillary 3D digital dental casts.	n = 16 (7M, 9F)	Median: 8.0 (range: 6.0–9.3) years	Growing	Orthodontic treatment: placement of resin modified glass ionomer cement on occlusal surfaces of selected lower teeth to treat dental anterior cross bite.	Median: 15.1 (range: 7.2–21.8) months	Dental models: surface-based	Dental models Area A: medial 2/3 of the third rugae and the area 5mm dorsal to them, Area B: superimposition on area A adding a 6mm wide stripe on the midpalate suture extending to the level of a line connecting the lingual grooves of the first permanent molars, Area C: superimposition on area A, but starting anteriorly from the medial 2/3 of the second rugae, Area D: superimposition on the palate bordered by a line 5mm from all gingival margins and extending to the middle of the first permanent molars, Area E: superimposition on an area of similar dimension as area A, but starting anteriorly at a line connecting the interproximal areas between the primary molars.
Yun et al Korean J Orthod (2018) (15)	Repeatability of serial intraoral scanner derived 3D model superimposition for evaluation of orthodontic tooth movement.	n = 7 (gender NA)	Mean: 22.0 ± 8.4 years	Growing and non-growing	Orthodontic treatment with fixed orthodontic appliances (extraction and non-extraction cases).	1 month (day of bonding: T0; 1 month later: T1)	Dental models: surface-based	Dental models: The palatal rugae and palatal slope separated by 5 mm from the gingival margins of the bilateral posterior teeth, and not extending distally beyond the line in contact with the distal surfaces of the bilateral first molars

NA: not available, M: male, F: female

Table 2. Main analysis-related characteristics and conclusions of the included studies.

Study name	Comparison/Control	Type of analysis	Method error	Conclusions	Limitations
An et al J Orofac Orthop (2015) (3)	Tooth movement assessed through surface-based superimposition of 3D dental models vs. Cephalometric radiographs.	Descriptive statistics only (medians, maximum, and all single measurements).	Intrarater on the whole process and sample (random error: Dahlberg's formula).	The tori seem to be more reliable for superimposition than the alveolar bone.	Only adult patients, sample size too small, comparison with cephalometric radiographs.
Becker et al J Orofac Orthop (2018) (6)	Tooth movement assessed through 10 landmarks superimposition vs. Surface-based superimposition of 3D dental models.	Descriptive statistics (mean, standard deviation), tests for normality (assessment for skewness, kurtosis and ShapiroWilk), linear regression models (normality of the residuals and homogeneity of variance were tested in advance).	NA	Landmark-based and an automated surface-matching approach may both allow for comparable results, though individual differences are evident.	Only mean values, no Bland-Altman plots, large age range, no method error, poor reporting.
Cha et al Eur J Orthod (2007) (10)	Tooth movement assessed through surface-based superimposition of 3D dental models vs. Cephalometric radiographs.	Paired t-tests, scattergrams and regression lines.	NA	There is no significant difference between the results of superimpositions of 3D digital models compared to those of cephalometric radiographs.	Comparison with cephalometric radiographs, no method error, no Bland Altman plots.
Chen et al Orthod Craniofac Res (2011) (11)	Tooth movement assessed through surface-based superimposition of 3D dental models vs. superimposed 3D dental models on stable miniscrews.	Descriptive statistics and pairwise comparisons for the main outcome (unpaired t-test) and intra-class correlation coefficient (ICC) for intra- and inter-observer error.	Intra and interrater.	The medial 2/3 of the third rugae and the palatal vault dorsal to it seem to be a stable superimposition reference in adult patients treated with premolar extractions.	Stable screws evaluated through linear measurements (0.5 mm allowance) and not through superimposition, deviation between structures visually assessed, results applicable only in non-growing patients, mean comparisons and not individual differences were assessed, the results of the validation/comparison were expected because the stable area was identified through the miniscrew superimposition.
Choi et al Angle Orthod (2010) (12)	Simulated tooth movement measurements made directly on plaster models vs. Surface-based superimposition of 3D dental models.	Paired t-test, Pearson correlation analysis (including scatter plots).	Intrarater on the whole process and sample.	There is no significant difference between measurements made directly on the plaster models and on superimposed digital 3D models.	In vitro study, no real orthodontic treatment, identical palatal vault and teeth.
Choi et al Korean J Orthod (2012) (13)	Tooth movement assessed through surface-based superimposition of 3D dental models vs. Cephalometric radiographs.	Bland-Altman plots using 95% limits of agreement and intra-class correlation (ICC).	Intrarater on the whole process and sample.	The 3D model superimposition is as clinically reliable for assessing antero-posterior tooth movement as cephalometric superimposition in cases treated by RME and maxillary protraction headgear. However, vertical tooth movements did not demonstrate adequate agreement.	Comparison with cephalometric radiographs.
Ganzer et al Eur J Orthod (2017) (14)	Tooth movement assessed through surface-based superimposition of 3D dental models vs. Original value of simulated tooth movement measurements.	Plot with the individual differences of each observer from the true value. Correlation statistics (mixed effects model with true value as a covariate) and intra-class correlation (ICC) for intra-observer error.	Intrarater and interrater on the whole process and sample.	The surface-based superimposition method seems to be valid when tested in duplicated, artificially changed 3D digital models.	No actual serial models, the palate was changed to simulate growth, but no details are provided.

Table 2. Continued

Study name	Comparison/Control	Type of analysis	Method error	Conclusions	Limitations
Jang et al Angle Orthod (2009) (7)	Tooth movement assessed through landmark-based followed by surface-based superimposition of 3D dental models vs. superimposed 3D dental models on stable miniscrews.	Pearson's correlation analysis and paired t-test.	Intraclass correlation.	The maxillary dental casts might be superimposed reliably using the medial points of the third rugae and the palatal vault.	Only mean values, no Bland-Altman plots, small sample size, no assessment of the validity of the gold standard method, treatment time not mentioned, no proper/detailed method error assessment.
Nalcaci et al Korean J Orthod (2015) (8)	Tooth movement assessed through landmark-based superimposition of 3D dental models vs. Cephalometric radiographs.	Friedman test, Cronbach's alpha (intrarater).	Intrarater on 10 sets of randomly selected measurements.	The study claimed no significant differences between the mean results of each superimposition method.	Comparison with cephalometric radiographs, comparison of mean values without showing/testing individual measurements.
Talaat et al Eur J Orthod (2017) (9)	Tooth movement assessed through landmark-based superimposition vs. surface-based superimposition of 3D dental models vs. surface based-superimposition of CBCT radiographs.	Intra-class correlation (ICC), analysis of variance (ANOVA).	Intrarater on the whole process and sample.	Regarding the results landmark-based superimposition seems to deliver similar results compared to the other methods.	Mean differences assessed, no Bland-Altman, cranial base superimposition compared to palate superimposition in growing patients.
Vasilakos et al Sci Rep (2017) (2)	Tooth movement assessed through surface-based superimposition of 3D dental models on the gold standard area A vs. Other 4 areas (B, C, D, E).	Permutational multivariante analysis of covariance (MANCOVA) with factorial fixed or mixed effects models, Wicoxon signed-rank test, Monte Carlo asymptotic p-value (PERMANOVA), Permutational analysis of multivariante dispersions (PERMDISP), Bonferroni correction, Bland-Altmann method.	Intrarater and interrater on the whole process and sample.	The medial part of the third rugae and a small area dorsal to it showed adequate accuracy for superimposing serial casts of growing patients. The other reference areas cannot be suggested.	Comparison to an assumed gold standard.
Yun et al Korean J Orthod (2018) (15)	NA	Intra-class correlation (ICC), random error (Dahlberg formula).	Intrarater on 7 patients.	3D superimposition of serial 3D dental models may provide repeatable evaluation of orthodontic tooth movement.	Measurements are based on landmark identification, no comparison, only mean values, no Bland-Altman plots, small sample size.

superimposition. There was no significant difference between the results of the techniques. This is one of the first studies in the topic and it has important limitations, such as the absence of method error, and the absence of individual measurement assessment. Ganzer *et al.* (14) evaluated a three-step surface-based superimposition technique (RFD: raw matching, fine matching, deformation analysis), based on simulated tooth movement and growth. This technique seems to be valid when tested in duplicated, artificially changed 3D digital models. However, further testing in actual patient data is required to confirm these findings. The palate was changed to simulate growth, but no details are provided on the way and the extent this was performed. Talaat *et al.* (9) focused on two surface-based methods, as well as on a landmark-based method. The landmark-based method will be discussed in the Superimposition on or adjacent to the incisive papilla section. The surface-based methods included the 3dMD superimposition of 3D dental models and the Invivo Dental superimposition of CBCTs. The study group included only growing patients and showed comparable results for all three software and methods. However, the study assessed only mean differences and a cranial base superimposition was compared to the palate superimposition in growing patients, which is questionable due to growth considerations.

Overall, all three studies have a high risk of bias, as well as high applicability concerns (9, 10, 14). Moreover, the heterogeneity in this subgroup is high, and thus, the available studies cannot be summarized to provide considerable amount of evidence for the use of the whole palate as superimposition reference.

The study of Jang *et al.* (7) evaluated a mixed superimposition method. The digital 3D dental models were first superimposed on the midpoint of the line connecting the medial points of the right and left third palatal rugae. In a second step, the definitive superimposition was done on a surface including a palatal vault area, but no rugae. Thus, this is a landmark-based superimposition followed by a surface-based superimposition. The study shows that the maxillary dental casts might be superimposed reliably on the described area, but it has high risk of bias and applicability concerns. There are several limitations as there were only mean values assessed, the sample size is very small, there is no assessment of the validity of the gold standard method, and there is no proper/detailed method error assessment. Thus, the present technique requires further validation.

The study of Becker *et al.* (6), which was also described in a previous category, compares two superimposition approaches. In the ICP surface-based superimposition approach the dental models are superimposed on a major region of the palate. This method shows comparable results with a landmark method, but the study is characterized as high risk of bias and has high applicability concerns.

Overall, both of the aforementioned studies show promising results (6, 7), but they have important limitations, and thus they were judged as high risk of bias and had high applicability concerns. Furthermore, the high heterogeneity of the studies regarding samples, methods, and outcomes did not allow for safe conclusions.

Superimposition on or adjacent to the incisive papilla

There are two studies that examined the superimposition on or adjacent to the incisive papilla and both included growing patients (8, 9). These studies have a high risk of bias, as well as high applicability concerns. One study that has also been reported in another category (9), used a landmark-based superimposition on one point at the incisive papilla and two points at the midpalatal suture. Among other limitations, the time interval between subsequent models was limited to 3 months, which is pretty short to generalize findings for regular orthodontic treatments. Thus, the validity of this technique remains questionable. In the second study that evaluates the superimposition on the incisive papilla (8), digital 3D models were superimposed on three points of the incisive papilla. The study claimed no statistically significant differences between the mean results of this superimposition method and cephalometric measurements. However, only mean values were showed/tested, not taking into account differences in individual measurements.

On the basis of the aforementioned studies (8, 9) no conclusion can be drawn for the use of the incisive papilla as a superimposition reference structure.

Discussion

In response to the rapid incorporation of intraoral scanners in everyday clinical practice and the huge possibilities they offer, we attempted in the present review to summarize and assess the available evidence on techniques for superimposition of serial 3D digital intraoral models. Such techniques are very useful to assess changes that occurred in a patient's mouth during time due to treatment, growth, or pathology. The superimposition of digital 3D models warrants a risk-free imaging technique, which is far more informative and with fewer limitations compared to the traditional superimposition of lateral cephalograms.

In total, a considerable amount of studies was included for analysis, but the high heterogeneity in hypotheses, samples, methods, and outcomes, as well as the high risk of bias and applicability concerns that characterized most studies did not allow for solid conclusions. Therefore, there is a strong need for further high-quality methodological studies, on different samples that will represent the broad spectrum of clinical conditions, in order to develop and validate these techniques. This will allow us to take full advantage of the big amount of high-quality information that the digital 3D intraoral models offer.

Since the 3D scanners were developed and used in the field in the last years we did not expect any study published previous to 2000. Indeed, though we did not apply any time restriction on our search, the earliest included study was published in 2007.

As mentioned in Materials and methods, a meta-analysis was planned if there were at least two studies of similar comparisons reporting the same outcomes at similar follow-up periods. Unfortunately, the heterogeneity of the included studies was too high, there was no consistent primary outcome, no similar samples or consistent interventions. Thus, a meta-analysis was not feasible.

We considered for eligibility all studies that deal with the identification of appropriate regions to superimpose digital 3D dental models. There are studies that are closely related to the topic, but are not included in this review. These studies are reported in the Supplementary Text, including the reason for exclusion. Finally, we included in our review a moderate amount of studies, but almost all of them examine different hypotheses on varying samples, which makes it hard to obtain comparable results.

Another problem is that the risk of bias of most studies is high. This could be attributed to the fact that the field of study is relatively new. There is no study with a low risk of bias and there are only two studies (2, 13) with low applicability concerns. Furthermore, there was no study with a robust gold standard technique to compare with. Various studies compared the 3D superimposition outcomes with cephalometric radiographs (3, 8, 10, 13). As already mentioned, the cephalometric radiographs have various inherent limitations and provide much less and lower quality information compared to the 3D intraoral models. Thus, such comparisons are probably inappropriate and the results of these studies should be treated with caution. In most studies the researchers mainly compared their technique against a sometimes questionable, self-assumed gold standard or against the inferior traditional method. The few studies that used implants as standard reference did not evaluate adequately their stability (7, 11).

There are different areas that may be reliable to superimpose maxillary digital 3D models. The area that includes the medial two-thirds of the third rugae and the region 5 mm dorsal to them might be a reliable reference. This area was reported as being stable through different studies (2, 7, 11), though most of them had high risk of bias. According to other studies, the superimposition on a major region of the palate might lead to acceptable superimposition results. This option was investigated by multiple studies (2, 6, 7, 9, 10, 12–15) and some of them tested exactly the same superimposition area. Namely, the *in vitro* study of Choi *et al.* (12) was confirmed by the clinical study of Choi *et al.* (13) and the study of Yun *et al.* (15). However, the aforementioned studies showed high heterogeneity. A limitation of superimposing on a major region of the palate is that it will probably include areas that are not stable. Thus, the limited available evidence also in this case emphasizes the need for further well-designed studies.

Although the superimposition on or adjacent to the incisive papilla is suggested by certain studies (8, 9), one should be sceptic about these methods. The incisive papilla is subjected to changes in shape and location due to various factors, such as growth, traumatic-occlusion or inflammation induced oedema (18). Moreover, the incisive papilla is very anteriorly located, and, therefore, small inaccuracies can exert a big effect in the outcomes at posterior regions.

The superimposition of digital 3D models in the mandible is reported in only one study (3) that could not identify a reliable reference area. The superimposition on the alveolar bone is also not adequately supported by the available studies so far. This is reasonable since extended remodelling of the alveolar bone is expected during growth and during orthodontic treatment (19, 20).

Finally, at the moment, there is no landmark-based method that is proved to deliver reliable results.

Limitations

A limitation of this review could be that the high heterogeneity among studies, combined with the quality assessment grades, did not allow for any adequate synthesis of the results.

Conclusion

3D digital models offer great opportunities for the assessment of tooth movement during time. So far, several techniques are recommended in the literature based on insufficient evidence. Until now, there is almost no evidence to support the superimposition of mandibular digital 3D models. However, there are multiple areas that might be appropriate for maxillary digital 3D model superimposition. According to this study the following areas of the maxilla could provide reliable outcomes, though not yet based on solid evidence:

1. The medial two-thirds of the third rugae and the area 5 mm dorsal to them.
2. An area including all rugae, with the lateral margins located at least 5 mm from the gingival margins and a distal margin that does not extend beyond the first molars.

From this study it is evident that there is an urgent need for further research in order to be able to properly incorporate these new, useful tools in research and in everyday clinical practice. Future studies should adequately test the performance of varying superimposition techniques in different participant groups, time spans, and treatments.

Supplementary material

Supplementary data are available at *European Journal of Orthodontics* online.

Funding

None.

Conflict of interest

None declared.

References

1. Halazonetis, D.J. (2005) From 2-dimensional cephalograms to 3-dimensional computed tomography scans. *American Journal of Orthodontics and Dentofacial Orthopedics,* 127, 627–637.
2. Vasilakos, G., Schilling, R., Halazonetis, D. and Gkantidis, N. (2017) Assessment of different techniques for 3D superimposition of serial digital maxillary dental casts on palatal structures. *Scientific Reports,* 7, 5838.
3. An, K., Jang, I., Choi, D.S., Jost-Brinkmann, P.G. and Cha, B.K. (2015) Identification of a stable reference area for superimposing mandibular digital models. *Journal of Orofacial Orthopedics,* 76, 508–519.
4. Cassetta, M., Altieri, F., Di Giorgio, R. and Silvestri, A. (2015) Two-dimensional and three-dimensional cephalometry using cone beam computed tomography scans. *Journal of Craniofacial Surgery*, 26, 311–315.
5. Gkantidis, N., Schauseil, M., Pazera, P., Zorkun, B., Katsaros, C. and Ludwig, B. (2015) Evaluation of 3-dimensional superimposition techniques on various skeletal structures of the head using surface models. *PLoS One,* 10, e0118810.
6. Becker, K., Wilmes, B., Grandjean, C., Vasudavan, S. and Drescher, D. (2018) Skeletally anchored mesialization of molars using digitized casts and two surface-matching approaches: analysis of treatment effects. *Journal of Orofacial Orthopedics,* 79, 11–18.
7. Jang, I., Tanaka, M., Koga, Y., Iijima, S., Yozgatian, J.H., Cha, B.K. and Yoshida, N. (2009) A novel method for the assessment of three-dimensional tooth movement during orthodontic treatment. *Angle Orthodontist,* 79, 447–453.
8. Nalcaci, R., Kocoglu-Altan, A.B., Bicakci, A.A., Ozturk, F. and Babacan, H. (2015) A reliable method for evaluating upper molar distalization: superimposition of three-dimensional digital models. *Korean Journal of Orthodontics,* 45, 82–88.
9. Talaat, S., Kaboudan, A., Bourauel, C., Ragy, N., Kula, K. and Ghoneima, A. (2017) Validity and reliability of three-dimensional palatal superimposition of digital dental models. *European Journal of Orthodontics,* 39, 365–370.
10. Cha, B.K., Lee, J.Y., Jost-Brinkmann, P.G. and Yoshida, N. (2007) Analysis of tooth movement in extraction cases using three-dimensional reverse engineering technology. *European Journal of Orthodontics,* 29, 325–331.
11. Chen, G., Chen, S., Zhang, X.Y., Jiang, R.P., Liu, Y., Shi, F.H. and Xu, T.M. (2011) Stable region for maxillary dental cast superimposition in adults, studied with the aid of stable miniscrews. *Orthodontics and Craniofacial Research,* 14, 70–79.
12. Choi, D.S., Jeong, Y.M., Jang, I., Jost-Brinkmann, P.G. and Cha, B.K. (2010) Accuracy and reliability of palatal superimposition of three-dimensional digital models. *Angle Orthodontist,* 80, 497–503.
13. Choi, J.I., Cha, B.K., Jost-Brinkmann, P.G., Choi, D.S. and Jang, I.S. (2012) Validity of palatal superimposition of 3-dimensional digital models

in cases treated with rapid maxillary expansion and maxillary protraction headgear. *Korean Journal of Orthodontics,* 42, 235–241.

14. Ganzer, N., Feldmann, I., Liv, P. and Bondemark, L. (2018) A novel method for superimposition and measurements on maxillary digital 3D models-studies on validity and reliability. *European Journal of Orthodontics,* 40, 45–51.
15. Yun, D., Choi, D.S., Jang, I. and Cha, B.K. (2018) Clinical application of an intraoral scanner for serial evaluation of orthodontic tooth movement: a preliminary study. *Korean Journal of Orthodontics,* 48, 262–267.
16. Whiting, P.F., Rutjes, A.W., Westwood, M.E., Mallett, S., Deeks, J.J., Reitsma, J.B., Leeflang, M.M., Sterne, J.A. and Bossuyt, P.M.; QUADAS-2 Group. (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine,* 155, 529–536.
17. McInnes, M.D.F., et al.; the PRISMA-DTA Group. (2018) Preferred Reporting Items for a Systematic Review and Meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA,* 319, 388–396.
18. Danz, J.C., Greuter, C., Sifakakis, I., Fayed, M., Pandis, N. and Katsaros, C. (2014) Stability and relapse after orthodontic treatment of deep bite cases-a long-term follow-up study. *European Journal of Orthodontics,* 36, 522–530.
19. Li, Y., Jacox, L.A., Little, S.H. and Ko, C.C. (2018) Orthodontic tooth movement: the biology and clinical implications. *Kaohsiung Journal of Medical Sciences,* 34, 207–214.
20. Björk, A. and Skieller, V. (1983) Normal and abnormal growth of the mandible. A synthesis of longitudinal cephalometric implant studies over a period of 25 years. *European Journal of Orthodontics,* 5, 1–46.